



Validation of TRMM Data with *In-Situ* Measurements over Colombia at Diurnal, Seasonal, Annual and Interannual Time Scales

Germán Poveda¹, Mauricio Bedoya¹, Sara M. Vallejo-Bernal¹, Alejandro Olarte¹, Viviana Urrea¹, Paniela Posada¹, Yadira Cárdenas-Posso², Franklyn Ruiz², María Teresa Martínez², Walter A. Petersen³ and George J. Huffman⁴.

e-mail: gpoveda@unal.edu.co

Department of Geosciences and Environment, Universidad Nacional de Colombia. ³Earth Sciences Office, NASA Marshall Space Flight Center, Huntsville, Alabama, USA. ⁴Mesoscale Atmospheric Processes Laboratory, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.



1. Introduction

Owing to its location in northwestern South America, Colombia exhibits complex geographical, geomorphological and hydro-climatological features arising from its tropical setting, but also from (Poveda et al., 2005): (i) the strong topographic gradients of the three ranges of the Andes crossing from southwest to northeast, (ii) atmospheric circulation patterns over the surrounding Pacific ocean and Caribbean Sea, (iii) the hydro-climatic dynamics of the Amazon and Orinoco River basins, (iv) the atmospheric circulation patterns over the neighboring tropical Pacific and Atlantic oceans, and (v) strong land-atmosphere feedbacks. In terms of the spatial distribution of precipitation, Snow (1976) refers to the Colombian Andes a "a dry island in a sea of rain". Along the Pacific coast, Colombia houses one of the rainiest places on Earth with mean annual rainfall between 12,000 and 13,000 mm per year. Such an extraordinary complexity provides an interesting setting to study the spatiotemporal dynamics of rainfall across a wide range of timescales, to which satellite information can further contribute. Here we perform a thorough validation study of TRMM rainfall data at a wide range of timescales using in-situ information from a large set of raingauges throughout Colombia. Using diverse metrics, validation is performed at diurnal, seasonal, annual and interannual timescales.

2. Objective

To develop a detailed validation study of different NASA satellite rainfall products over Colombia such as TRMM 3B42 and TRMM 3B43 from the Precipitation Measurement Missions, using ground precipitation records from a large set of raingauges, at different temporal resolutions (diurnal, monthly, seasonal, annual and interannual) and during different phases of diverse macroclimatic events including ENSO, PDO, NAO and AMO.

3. Data

3.1. Rain gauges

We used rainfall data gathered by Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM) for diverse periods and temporal resolutions, as follows: 260 raingauges with hourly records spanning from August 1988 to January 2014, and 2,683 raingauges stations with daily and monthly records from November 1900 to November 2016. From this data set, we selected 223 raingauges to perform the validation at diurnal timescale from 1998 to 2013. For validation at daily timescale, we selected 1,145 raingauges, those with maximum 10% of missing data during the period 1998-2014. Towards the validation at seasonal, annual and interannual timescales, we selected and reconstructed time series of precipitation from 1,662 raingauges, those exhibiting a maximum of 30% of missing data during the validation period 1998-2015

3.2. TRMM 3B42 v7 and TRMM 3B43 v7

We used the NASA satellite products TRMM 3B42 and TRMM 3B43, both with regular spatial resolution of 0.25°. For diurnal validation, we used 3-hourly TRMM 3B42 from 1998 to 2013 5.4. Interannual validation of TRMM 3B43 v7 and selected 142 pixels, those with at least one station of IDEAM; if the pixel contains two or more raingauges inside, we averaged rainfall data. For daily validation, we used daily TRMM 3B42 from 1998 to 2014 and selected 519 pixels, in this case if the pixel has two or more raingauges, one of those was selected randomly. For seasonal, annual and interannual validation, we used monthly TRMM 3B43 from 1998 to 2015 and selected 632 pixels, if the pixel contained two or more raingauges inside, we averaged their respective rainfall data.

4. Methodology

To validate TRMM at seasonal, annual and interannual scales, we first reconstruct monthly time series using the random orthogonal model and convex optimization [Candès & Recht, 2009]. We retain the inter-annual variability filtering out the annual cycle of precipitation from TRMM and IDEAM datasets through a monthly standardization procedure. We used Singular Value Decomposition (SVD) to identify the main variability modes from each dataset. SVD is useful to identify spatial patterns of those variability modes, and their temporal dynamics are called Singular Vectors. To perform the conditional analysis included in the study of interannual variability, we combine the ENSO phases (El Niño and La Niña) with the positive and negative phases of macro-climatic phenomena such as AMO, NAO and PDO. For diurnal and daily validations, we use the original series without reconstructing missing data. We use diverse metrics for validation purposes, including the Kolmogorov-Smirnov test of equality between Probability Distributions Functions (PDF), and diverse standard metrics such as absolute and relative errors and bias rates. A detailed description of these metrics is presented in Table 1.

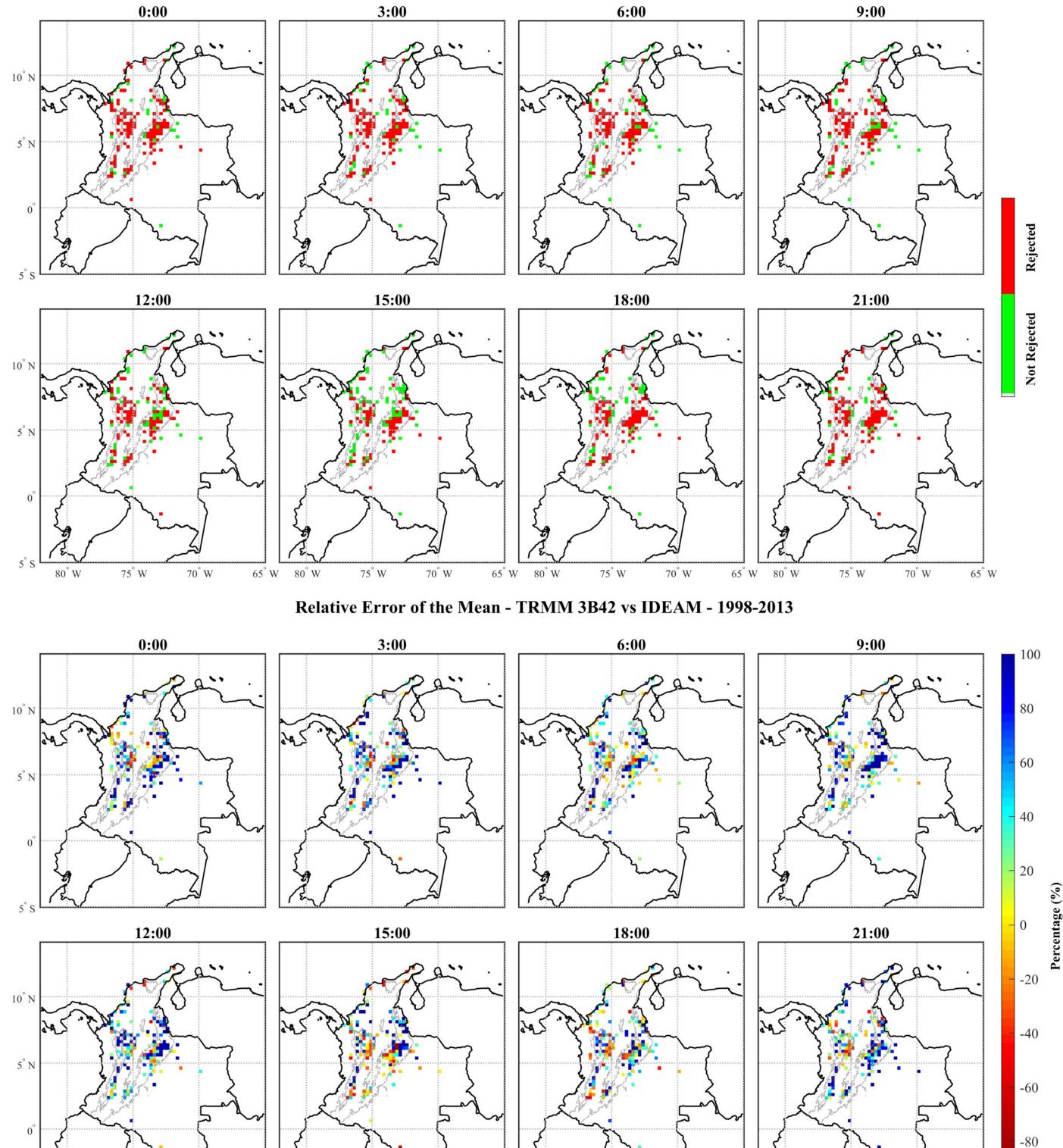
Table 1. Error metrics used for validation

Statistic	Equation	Units	Range	Perfect score
Mean absolute error	$MAE = \frac{1}{N} \sum TRMM - IDEAM $	mm	$[0,\infty)$	0
Relative error	$RE = \frac{\frac{1}{N} \sum (TRMM - IDEAM)}{IDEAM} *100$	%	$(-\infty, \infty)$	0
Kolmogorov- Smirnov	$\mathbf{D}_{KS} = \max_{X} \left \hat{F}_{TRMM}(x) - F_{IDEAM}(x) \right $	Dimensionless	[0,1]	0
Bias rate	$BIAS = \frac{\sum TRMM}{\sum IDEAM}$	Dimensionless	$[0,\infty)$	1

5. Results

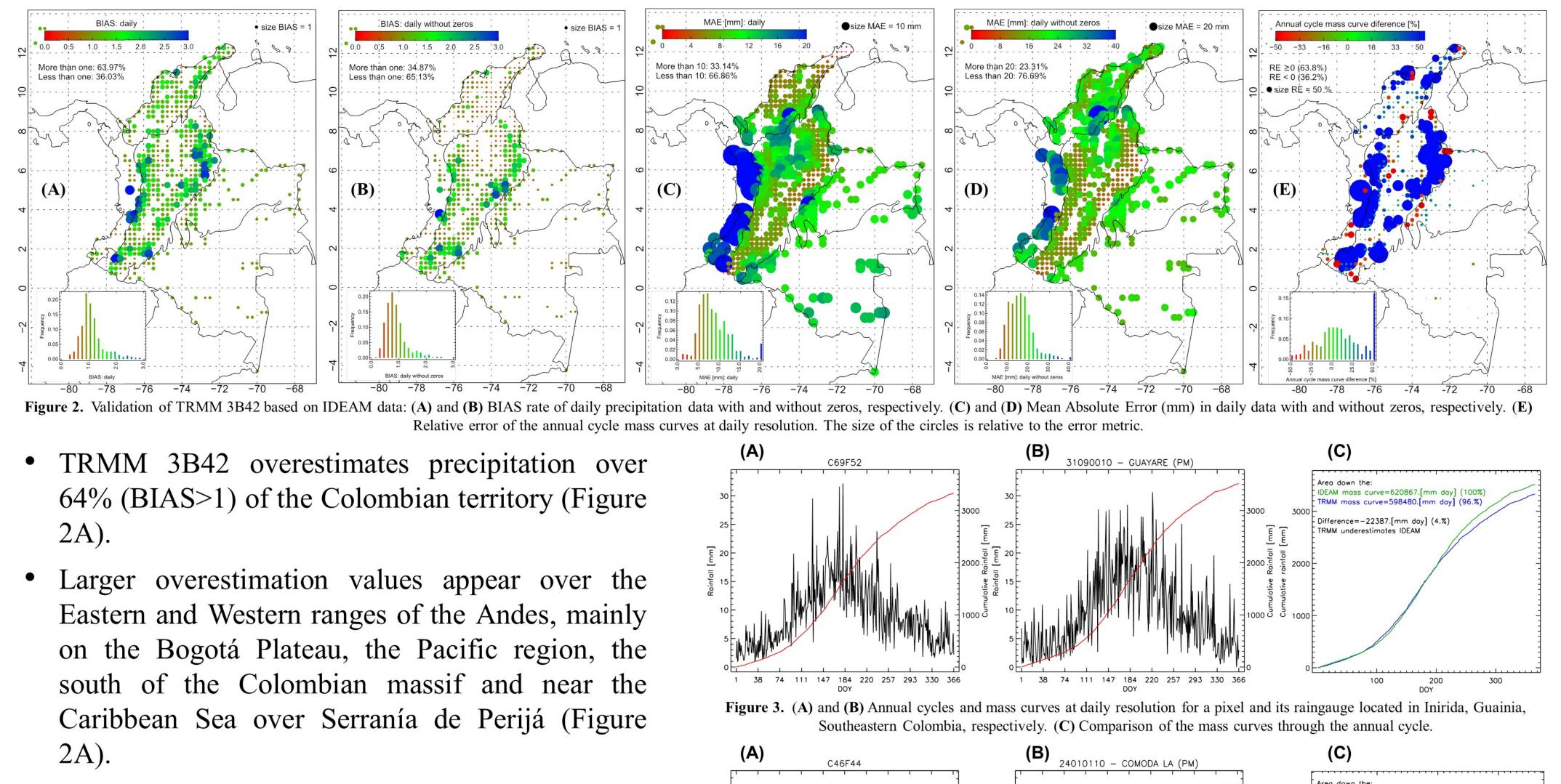
5.1. Diurnal validation of TRMM 3B42 v7

Kolmogorov-Smirnov h_o - TRMM 3B42 vs IDEAM - 1998-2013



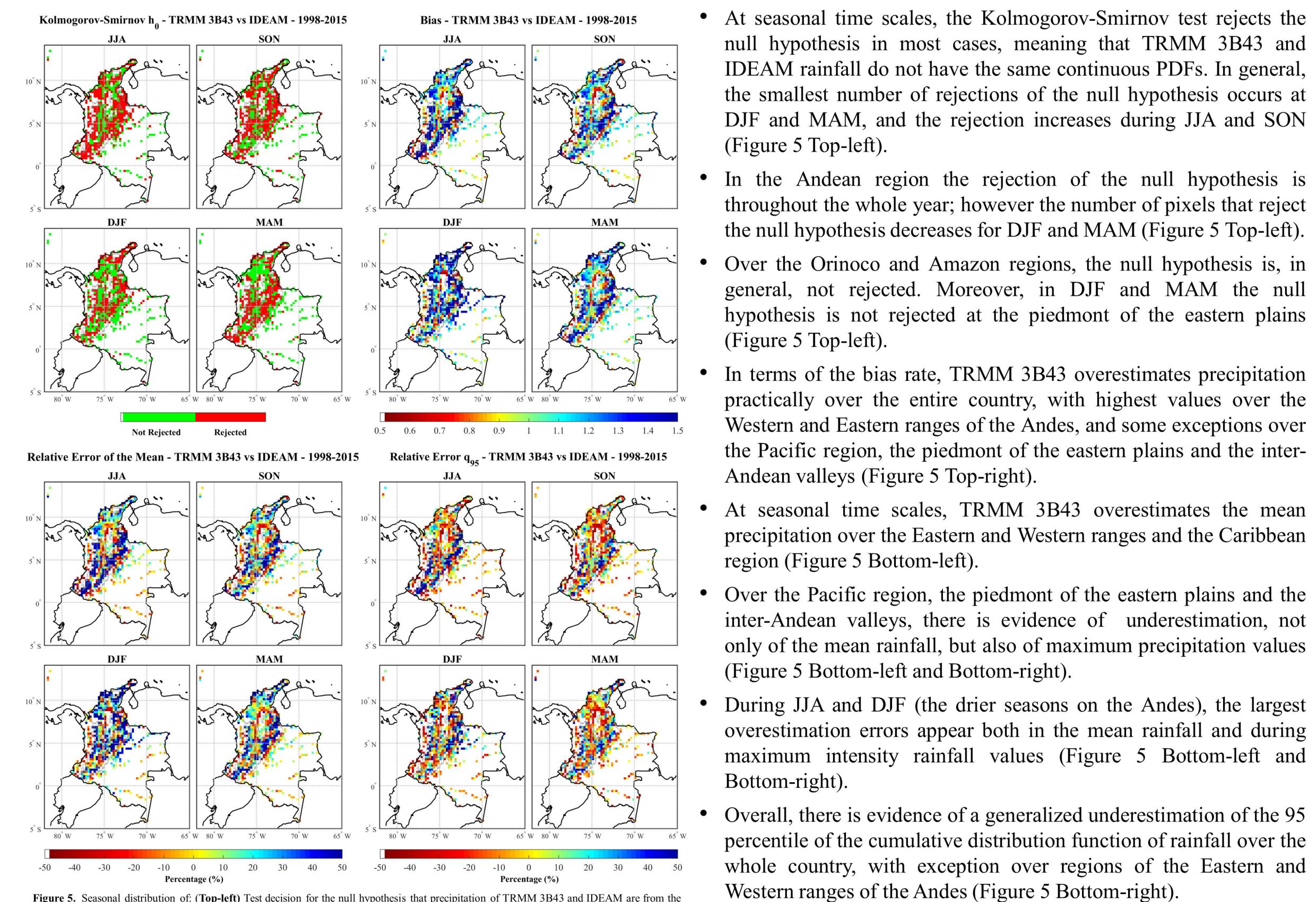
- Overall, TRMM 3B42 overestimates the mean hourly rainfall over the Eastern and Western ranges of the Andes (Figure Bottom).
- At diurnal scale, the Kolmogorov-Smirnov test rejects the null hypothesis in most cases, which means that precipitation from TRMM 3B42 and IDEAM do not have the same continuous PDF (Figure 1 Top).
- Over the Andes, the rejection of the null hypothesis is persistent for any 3-hour period of the day, however the number of pixels that reject the null hypothesis decreases from 12:00 to 18:00 hours (Figure 1 Top).
- In the Caribbean region the rejection of the null hypothesis prevails from 18:00 to 00:00 hours. In the Orinoco and Amazon regions, the lowest number of rejections occurs between 03:00 and 9:00 hours (Figure 1 Top).
- In general, the smallest number of rejections of the null hypothesis occurs at 15:00 hours (Figure 1 Top).
- Over the entire Colombian territory, the largest overestimation occurs between 9:00 and 12:00 hours. The largest overestimation at any time of the day occurs over the Eastern cordillera along the Bogotá plateau (Figure 1
- The worst underestimation (-50% to -100%) occurs on the piedmonts of the Andean ranges. A dipole with overestimates and underestimates is formed along the canyon of the Cauca River, bordering the Nudo del Paramillo, approximately at 7°N-75.5W°, 3730 m a.s.l. (Figure 1 Bottom).

5.2. Daily and annual validation of TRMM 3B42 v7



- Without zeros, the underestimation reaches to 65% (BIAS<1) and occurs mainly along the Caribbean plains, Amazon and Orinoco regions and over the Magdalena River valley (Figure 2B).
- Similar overestimation patterns are observed on the Eastern and Western Andes at daily time scale without zeroes (Figure 2B).
- The highest MAE values are observed over the Pacific region, followed by the Andean region and the Eastern Llanos piedmont. In La Guajira, the northernmost tip of Colombia, the smaller MAE values are observed. However when analyzing the time series without zeroes, the entire Caribbean region presents significant errors, and the Eastern and Western ranges of the Andes exhibit the smallest MAE values in the whole country (Figure 2C-D).
- The relative error of the annual cycle mass curves at daily resolution reveals a general overestimation of TRMM 3B42 over 64% (RE>0%) of Colombia, mainly over the Eastern and Western ranges of the Andes, on the Bogotá Plateau, the Pacific region, the south of the Colombian massif, and near the the Caribbean over Serranía de Perijá (Figure 2E).

5.3. Seasonal validation of TRMM 3B43 v7

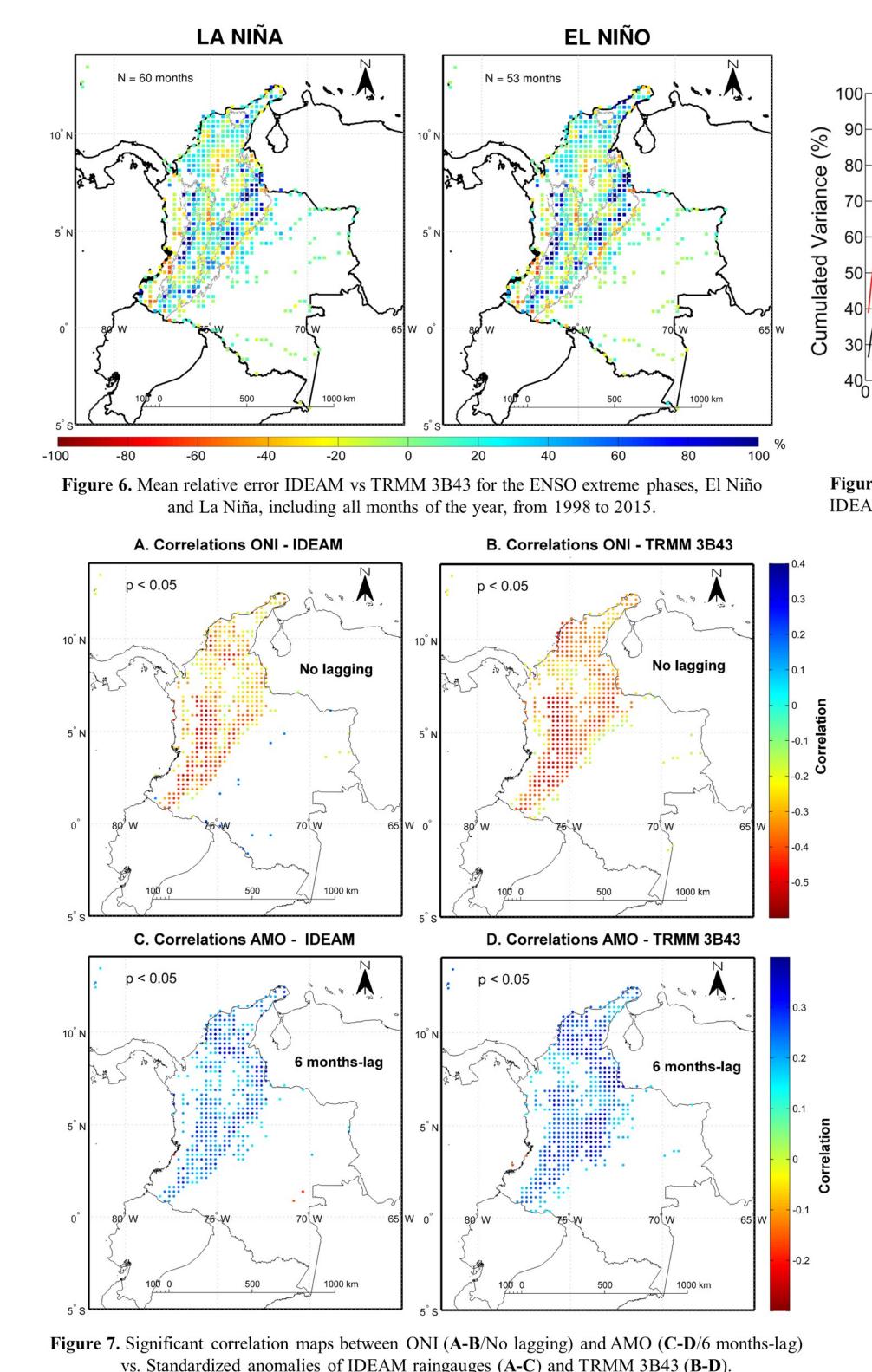


the mean precipitation. (Bottom-right) Relative error of the 95 percentile of the distribution function.

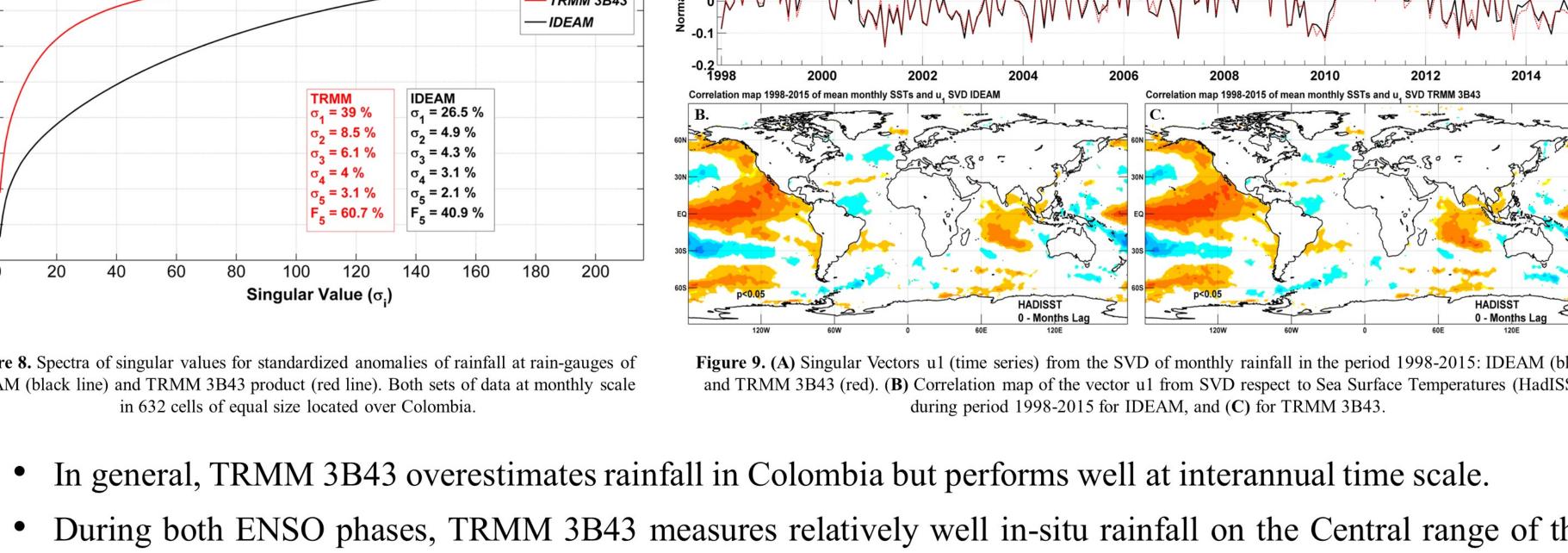
- DJF and MAM, and the rejection increases during JJA and SON (Figure 5 Top-left). In the Andean region the rejection of the null hypothesis is throughout the whole year; however the number of pixels that reject
- the null hypothesis decreases for DJF and MAM (Figure 5 Top-left). Over the Orinoco and Amazon regions, the null hypothesis is, in
- general, not rejected. Moreover, in DJF and MAM the null hypothesis is not rejected at the piedmont of the eastern plains
- In terms of the bias rate, TRMM 3B43 overestimates precipitation practically over the entire country, with highest values over the Western and Eastern ranges of the Andes, and some exceptions over the Pacific region, the piedmont of the eastern plains and the inter-Andean valleys (Figure 5 Top-right).
- At seasonal time scales, TRMM 3B43 overestimates the mean precipitation over the Eastern and Western ranges and the Caribbean region (Figure 5 Bottom-left).
- Over the Pacific region, the piedmont of the eastern plains and the inter-Andean valleys, there is evidence of underestimation, not only of the mean rainfall, but also of maximum precipitation values (Figure 5 Bottom-left and Bottom-right).
- During JJA and DJF (the drier seasons on the Andes), the largest overestimation errors appear both in the mean rainfall and during maximum intensity rainfall values (Figure 5 Bottom-left and
- Overall, there is evidence of a generalized underestimation of the 95 percentile of the cumulative distribution function of rainfall over the whole country, with exception over regions of the Eastern and Western ranges of the Andes (Figure 5 Bottom-right).
- as distribution function, using the two-sample Kolmogorov-Smirnov test. (Top-right) Bias rate. (Bottom-left) Relative error of • The seasons in which we observe greater underestimations of maximum rainfall are the wettest on the Andes: SON and MAM (Figure 5 Bottom-right).

Figure 1. Diurnal distribution of: (Top) Test decision for the null hypothesis that rainfall of TRMM 3B42 and IDEAM are from the same continuous

distribution function, using the two-sample Kolmogorov-Smirnov test. (Bottom) Relative error of the mean precipitation



Excellent representation of first interannual variability mode



- During both ENSO phases, TRMM 3B43 measures relatively well in-situ rainfall on the Central range of the Andes, but overestimates it over the Western and Eastern rages (Figure 6). TRMM 3B43 better represents in-
- TRMM 3B43 spectrum contains most of the variance in the first singular values, while IDEAM spectrum distributes it among the entire set of singular values (Figure 8). Otherwise, TRMM 3B43 optimally represents the main variability mode of raingauges data (Figure 9A-C).

situ rainfall on the low-lands such as the Caribbean, Orinoco and Amazon regions of Colombia.

the spatial pattern of relative error between TRMM and IDEAM over Colombia (Figure 6).

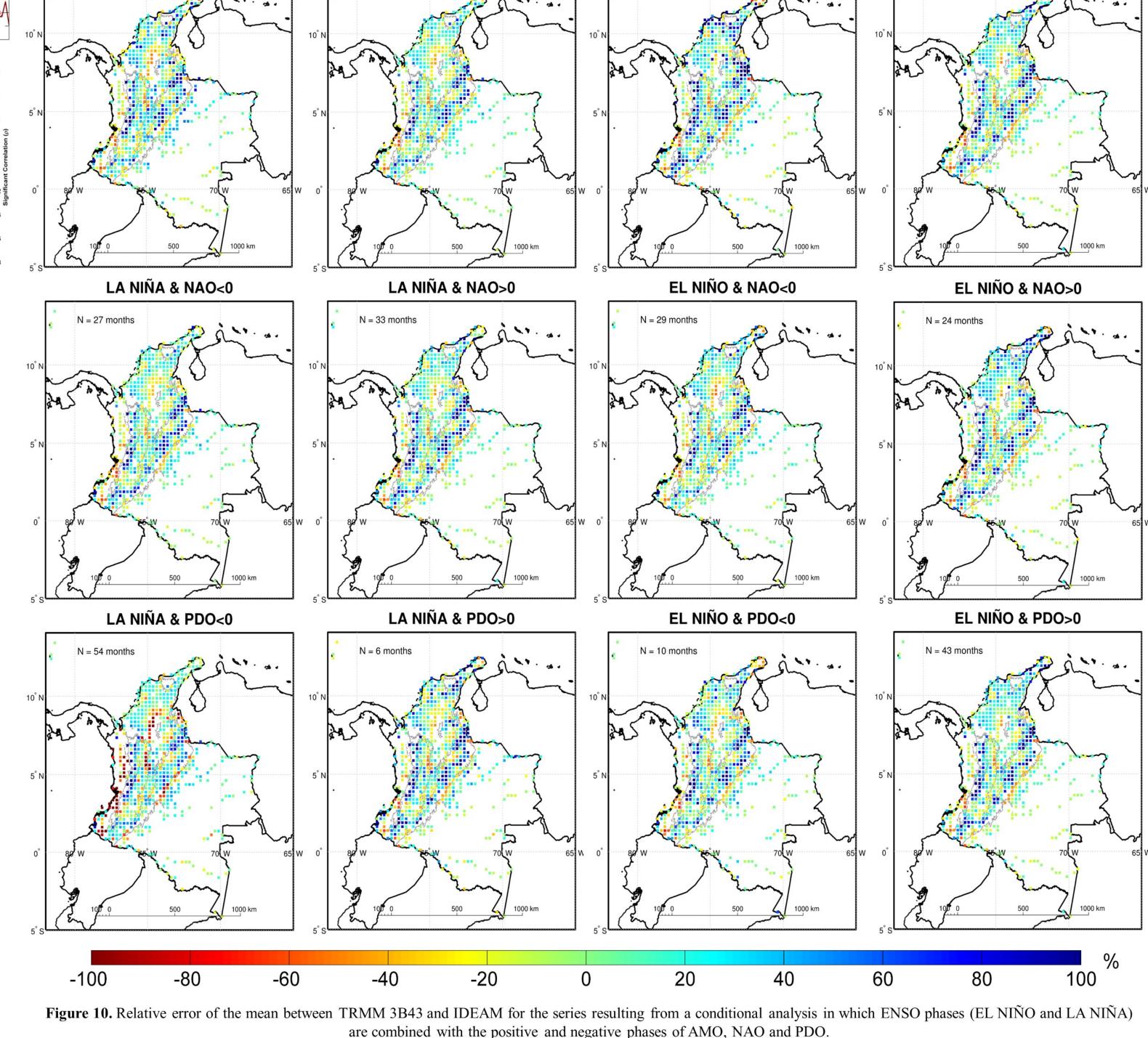
• TRMM 3B43 describes well the spatial pattern of correlations between ONI and in-situ rainfall anomalies (Figure 7A-B). TRMM 3B43 also describes quite well that AMO exhibits larger correlations with in-situ rainfall at 6 month-lags (Figure 7C-D). • ENSO dominates the interannual variability of rainfall in Colombia despite the fact that it is in phase with

other macro-climatic phenomena (Figure 10). It is clear that the ENSO phase (El Niño or La Niña) determines

- Among El Niño combinations with other macroclimatic indices (AMO-NAO-PDO), the cold phase of the AMO intensifies the TRMM 3B43 overestimation of rainfall. Beware the small data sample.
- Among La Niña combinations, the warm phase of the PDO intensifies the TRMM 3B43 overestimation of

In general, TRMM 3B43 overestimates rainfall in Colombia, but performs well at

the Eastern range of the Andes, respectively. (C) Comparison of the mass curves through the annual cycle.



6. Conclusions

- The diurnal cycle of precipitation at the Colombian Andes has its minimum al the tri-hour period centered at 09:00 hours (Poveda et al., 2005). That explains why one of the peaks of overestimation of the TRMM3B42 measurements occurs during that period. On the other hand, most underestimation case at daily timescale can be explained by a generalized underestimation at all tri-hour time periods (Figure 1 Bottom).
- At daily scale, general overestimation of rainfall values appear over the Eastern and Western ranges of the Andes, mainly on the Bogotá Plateau, the Pacific region, the south of the Colombian massif and near the Caribbean over Serranía de Perijá, and underestimation occurs mainly on the Caribbean, Amazon and Orinoco regions, and over Magdalena River valley (Figure 2A-E).
- At seasonal scale, there is a persistent overestimation of the mean and maximum precipitation values, over the Eastern and Western ranges of the Andes and the Caribbean region, and underestimation over the Pacific region, the piedmont of the eastern plains and the inter-Andean valleys. The errors of overestimation are exacerbated during JJA and DJF, the driest seasons on the Andes, while underestimation errors are exacerbated during SON and MAM, the wettest seasons on the Andes (Figure 5 Bottom-left and Bottom-right).
- At interannual scales, the estimated relative errors during El Niño and La Niña years indicate that TRMM 3B43 generally overestimates over less rainy regions (Bogotá Plateau, Eastern Andes, La Guajira and over the Magdalena River Valley, and the Patía River basin), and at the same time TRMM 3B43 underestimates it over the rainiest regions (Pacific Coast, northern Antioquia, and Catatumbo River basin). The relative magnitude and percentage of the error is larger during La Niña, and smaller during El Niño (Figure 6).
- The spatial patterns of concurrent correlations ONI-TRMM 3B43 and ONI-IDEAM are highly coherent over the Andean ranges, the Magdalena-Cauca River basins, and over the Caribbean low-lands (Figure 7A-B). In both datasets, correlations are significant and reverse (negative) on those regions. These results agree well with the reverse relationship between SSTs in the Pacific Equatorial and rainfall over the Andes of Colombia (Figure 9B-C). In both datasets (TRMM and IDEAM), the AMO spatial correlation patterns peaking at 6 month-lags are relevant over the Colombian Andes and the Caribbean low-lands.
- The relative error of the mean between TRMM 3B43 and IDEAM for the conditional analysis at interannual time scale, shows spatially, for all combinations, that in the Caribbean region and in the Eastern and Western ranges of the Andes there is an overestimation of rainfall, while in the Pacific region there is an underestimation of rainfall. Over the Amazon and Orinoco regions, TRMM 3B43 does an excellent job in capturing the main interannual variability mode associated with ENSO (Figure 10).

7. References

- Candès, E., & B. Recht (2009). Exact matrix completion via convex optimization. Communications of the ACM, 55(6), 111-119. Poveda, G., O. J. Mesa, L. F. Salazar, and co-authors, (2005). The diurnal cycle of precipitation in the Tropical Andes of Colombia. Mon. Wea. Rev., 133, 228-240. doi: http://dx.doi.org/10.1175/MWR-2853.1
- Snow, J. W., (1976). The climate of northern South America, In: Climates of Central and South America, W. Schwerdtfeger (ed.), Elsevier, Amsterdam, 295-403.

Acknowledgments

We gratefully acknowledge the financial support of Universidad Nacional de Colombia through the Program "Convocatoria Nacional de Proyectos" para el Fortalecimiento de la Investigación, Creación e Innovación de la Universidad Nacional de Colombia 2016-2018".